Some Mathematical And Numerical Questions Connected With First And Second Order Time Dependent Systems Of Partial Differential Equations

Heinz-O. Kreiss[1] and Omar E. Ortiz[2]

[1] Department of Mathematics, University of California, Los Angeles, CA 90095, USA
[2] Universidad Nacional de Córdoba, Facultad de Matemática, Astronomía y Física, (5000) Córdoba, Argentina.

1 Introduction

There is a tendency to write the equations of general relativity as a first order symmetric system of time dependent partial differential equations. However, for numerical reasons, it might be advantageous to use a second order formulation like one obtained from the ADM equations. Unfortunately, the type of the ADM equations is not well understood and therefore we shall discuss, in the next section, the concept of wellposedness. We have to distinguish between weakly and strongly hyperbolic systems. Strongly hyperbolic systems are well behaved even if we add lower order terms. In contrast; for every weakly hyperbolic system we can find lower order terms which make the problem totally illposed. Thus, for weakly hyperbolic systems, there is only a restricted class of lower order perturbations which do not destroy the wellposedness. To identify that class can be very difficult, especially for nonlinear perturbations. In Section 3 we will show that the ADM equations, linearized around flat with constant lapse function and shift vector, are only weakly hyperbolic. However, we can use the trace of the metric as a lapse function to make the equations into a strongly second order hyperbolic system.

Using simple models we shall, in Section 4, demonstrate that approximations of second order equations have better accuracy properties than the corresponding approximations of first order equations. Also, we avoid spurious waves which travel against the characteristic direction.

In the last section we discuss some difficulties connected with the preservation of constraints.

2 Well Posed Problems

2.1 First Order Systems

Consider the Cauchy problem for a first order system with constant coefficients

$$\boldsymbol{u}_t = \sum_{j=1}^{s} A_j D_j \boldsymbol{u} =: P(D)\boldsymbol{u}, \qquad D_j = \frac{\partial}{\partial x_j}, \tag{1}$$

$$\boldsymbol{u}(t=0)=\boldsymbol{f}(x), \qquad x=(x_1,\dots,x_s),\ -\infty<x_j<\infty.$$

We construct simple wave solutions

$$\boldsymbol{u}(x,t)=e^{\mathrm{i}(\omega,x)}\widehat{\boldsymbol{u}}(\omega,t), \quad \omega=(\omega_1,\dots,\omega_s) \ \text{ real}$$

and obtain

$$\widehat{\boldsymbol{u}}_t=\mathrm{i}|\omega|\sum_{j=1}^{s}A_j\omega_j'\widehat{\boldsymbol{u}}=:\mathrm{i}|\omega|P(\omega')\widehat{\boldsymbol{u}}, \qquad \omega'=\omega/|\omega| \tag{2}$$

Definition 1. We call the problem (1) strongly hyperbolic if for every ω' the eigenvalues of the symbol $P(\omega')$ are real and there is a complete set of uniformly (in ω') linearly independent eigenvectors.

Examples of strongly hyperbolic systems are symmetric systems where $A_j=A_j^{\dagger}$.

The solutions of strongly hyperbolic systems satisfy an energy estimate

$$\|\boldsymbol{u}(\cdot,t)\|^2\le K^2\mathrm{e}^{2\alpha t}\|\boldsymbol{u}(\cdot,0)\|^2. \tag{3}$$

Here K,α are universal constants which do not depend on the initial data $\boldsymbol{u}(x,0)=\boldsymbol{f}(x)$. The norms are L_2 norms. For systems (1) with constant coefficients $\alpha=0$.

Strong hyperbolicity and the existence of an energy estimate are equivalent, we have[1]

Theorem 1. *The solutions of* (1) *satisfy an energy estimate of the type* (3) *if and only if the system is strongly hyperbolic.*

The most important property of strongly hyperbolic systems is that we can add lower order terms and an estimate of type (3) is still valid. We have

Theorem 2. *Let* (1) *be strongly hyperbolic. Then the solutions of*

$$\boldsymbol{w}_t=P(D)\boldsymbol{w}+B\boldsymbol{w} \tag{4}$$

satisfy an estimate of type (3)*. Here B is any bounded operator.*

Definition 2. We call the problem (1) weakly hyperbolic if the eigenvalues of $P(\omega')$ are real.

In this definition we do not require that there is a complete set of eigenvectors. An example in dimension one is given by

$$\boldsymbol{u}_t=\begin{pmatrix}1&1\\0&1\end{pmatrix}\boldsymbol{u}_x=:A\boldsymbol{u}_x$$

[1] First order theory is well known, we refer to [1].

Simple wave solutions for this system have the form

$$\boldsymbol{u}(x,t) = \mathrm{e}^{\mathrm{i}\omega A t}\,\mathrm{e}^{\mathrm{i}\omega x}\widehat{\boldsymbol{u}}(\omega,0) = \left(I + \mathrm{i}\,\omega \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} t\right) \mathrm{e}^{\mathrm{i}\omega(x+t)}\widehat{\boldsymbol{u}}(\omega,0)$$

Thus there is no exponential growth but there is the polynomial growth in ωt. This is typically the case for weakly hyperbolic systems. One can prove

Theorem 3. *For weakly hyperbolic systems the growth of simple wave solutions is at most of the order $\mathcal{O}(1+|\omega t|^{n-1})$, where n is the number of components of $\boldsymbol{u}$.*

The real difficulty with weakly hyperbolic systems is that lower order terms can make them exponentially ill posed. For example, consider

$$\boldsymbol{u}_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \boldsymbol{u}_x + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \boldsymbol{u}.$$

Making the simple wave ansatz

$$\boldsymbol{u}(x,t) = \mathrm{e}^{\mathrm{i}\omega x}\widehat{\boldsymbol{u}}(\omega,t)$$

we obtain

$$\widehat{\boldsymbol{u}}_t = \begin{pmatrix} \mathrm{i}\omega & \mathrm{i}\omega \\ 1 & \mathrm{i}\omega \end{pmatrix} \widehat{\boldsymbol{u}} =: A\widehat{\boldsymbol{u}}.$$

The eigenvalues λ of A are given by

$$\lambda = \mathrm{i}\omega \pm \sqrt{\mathrm{i}\omega}, \quad \text{i.e. } \mathrm{Re}\lambda = \pm\frac{\sqrt{2}}{2}\sqrt{|\omega|}.$$

Therefore the perturbed problem is exponentially ill posed.

The same result holds if we consider the variable coefficient problem

$$\boldsymbol{u}_t = U^{\dagger}(x) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} U(x)\boldsymbol{u}_x, \quad U(x) = \begin{pmatrix} \cos x & \sin x \\ \sin x & \cos x \end{pmatrix}.$$

There are no problems to generalize the results to variable coefficients and quasi-linear systems if the system is pointwise strongly hyperbolic.

Unfortunately, in applications one can be confronted with systems which are weakly hyperbolic. In this case one has to carefully study the influence of lower order terms. For example, trivially,

$$u_t + (u^2)_x + (v^2)_x = -\alpha u, \quad v_t + (v^2)_x = -\alpha v$$

is well behaved ($\alpha > 0$ sufficiently large so that no shocks form). We can solve the second equation to obtain v which becomes a governing function in the first equation.

2.2 Second Order Systems

We consider second order systems

$$\boldsymbol{u}_{tt} = P_0(D)\boldsymbol{u} + P_1(D)\boldsymbol{u}_t \tag{5}$$

where

$$P_0(D) = \sum_{j,k=1}^{s} A_{jk} D_j D_k\,, \quad P_1(D) = \sum_{j=1}^{s} A_j D_j\,.$$

We calculate simple wave solutions. Introducing

$$\boldsymbol{u}(x,t) = e^{\mathrm{i}(\omega,x)}\widehat{\boldsymbol{u}}(\omega,t)$$

gives us

$$\widehat{\boldsymbol{u}}_{tt} = -|\omega|^2 P_0(\omega')\widehat{\boldsymbol{u}} + \mathrm{i}|\omega| P_1(\omega')\widehat{\boldsymbol{u}}_t. \tag{6}$$

We have

Lemma 1. *A necessary condition for well posedness is that, for all ω', the eigenvalues $\tilde{\kappa}$ of*

$$[-\tilde{\kappa}^2 I + P_1(\omega')\tilde{\kappa} + P_0(\omega')]\boldsymbol{a} = 0 \tag{7}$$

are real.

Proof. If $\tilde{\kappa}(\omega'), \boldsymbol{a}(\omega')$ is a solution of (7) then $-\tilde{\kappa}(\omega'), \boldsymbol{a}(\omega')$ is a solution if we replace ω' by $-\omega'$. Since the solutions of (6) are of the form $e^{\mathrm{i}|\omega|\tilde{\kappa}(\omega')t}\boldsymbol{a}(\omega')$ we only avoid catastrophic growth if $\tilde{\kappa}$ is real.

If $P_1 = 0$ then (7) becomes

$$[-\tilde{\kappa}^2 I + P_0(\omega')]\boldsymbol{a} = 0$$

and Lemma 1 reduces to

Lemma 2. *If $P_1 = 0$ then a necessary condition for well posedness is that the eigenvalues of $P_0(\omega')$ are real and nonnegative.*

We can write (6) as a first order system by introducing a new variable $\widehat{\boldsymbol{v}}$ with $\widehat{\boldsymbol{u}}_t = \mathrm{i}|\omega|\widehat{\boldsymbol{v}}$. We obtain

$$\begin{pmatrix} \widehat{\boldsymbol{v}} \\ \widehat{\boldsymbol{u}} \end{pmatrix}_t = \mathrm{i}|\omega| \begin{pmatrix} P_1(\omega') & P_0(\omega') \\ I & 0 \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{v}} \\ \widehat{\boldsymbol{u}} \end{pmatrix} =: \mathrm{i}|\omega|\widehat{\mathbf{P}} \begin{pmatrix} \widehat{\boldsymbol{v}} \\ \widehat{\boldsymbol{u}} \end{pmatrix} \tag{8}$$

The eigenvalues of $\widehat{\mathbf{P}}$ are determined by (7). Using this reduction, we can define what we mean by strongly and weakly hyperbolic (second order) systems.

Definition 3. The system (5) is called strongly hyperbolic if for all ω' the eigenvalues of $\widehat{\mathbf{P}}$ are real and there is a uniformly linearly independent (in ω') complete set of eigenvectors.

For strongly hyperbolic systems one can again develop a rather complete theory for local existence of quasi-linear systems. In particular lower order terms

$$Q\boldsymbol{u} = \sum_{j=1}^{s} B_j D_j \boldsymbol{u} + B_0 \boldsymbol{u}_t + C\boldsymbol{u}$$

do not destroy the well posedness of the problem.

If $P_1 \equiv 0$ we have

Theorem 4. *Assume that $P_1 \equiv 0$. The system is strongly hyperbolic if and only if the eigenvalues of $P_0(\omega')$ are strictly positive and $P_0(\omega')$ has a complete set of eigenvectors which is uniformly (in ω') linearly independent.*

Proof. Notice that when $P_1 \equiv 0$, any eigenvector of $\widehat{\mathbf{P}}$ with eigenvalue $\lambda_j(\omega')$ is of the form

$$\begin{pmatrix} \lambda_j \boldsymbol{a}_j \\ \boldsymbol{a}_j \end{pmatrix} \tag{9}$$

where the splitting corresponds to the block structure of $\widehat{\mathbf{P}}$. Moreover, for each eigenvector of the form (9) there is another eigenvector

$$\begin{pmatrix} -\lambda_j \boldsymbol{a}_j \\ \boldsymbol{a}_j \end{pmatrix}$$

with eigenvalue $-\lambda_j(\omega')$, which is linearly independent from the first if and only if $\lambda_j \neq 0$.

Now, it is easy to check that a set of eigenvectors

$$\left\{ \begin{pmatrix} \lambda_j \boldsymbol{a}_j \\ \boldsymbol{a}_j \end{pmatrix}; \begin{pmatrix} -\lambda_j \boldsymbol{a}_j \\ \boldsymbol{a}_j \end{pmatrix} \right\}, \qquad j = 1, 2, \ldots n.$$

with real eigenvalues $\{\lambda_j, -\lambda_j\}$, $j = 1, 2, \ldots, n$ is a set of $2n$ uniformly linearly independent (in ω') eigenvectors if and only if the set $\{\boldsymbol{a}_j(\omega')\}$, $j = 1, 2, \ldots n$, is a set of uniformly linearly independent (in ω') eigenvectors of $P_0(\omega')$ with positive eigenvalues $\lambda_j^2(\omega') > 0$. This proves the theorem.

Definition 4. The system (5) is called weakly hyperbolic if for all ω' the eigenvalues of $\widehat{\mathbf{P}}$ are real.

In particular we have

Lemma 3. *If $P_1 \equiv 0$ and $P_0(\omega')$ has zero as an eigenvalue then the system is not strongly hyperbolic. It is weakly hyperbolic if the eigenvalues are real and non negative.*

As in the case of first order systems, weakly hyperbolic systems can have catastrophic exponential growth when adding lower order terms or considering variable coefficients. As example we consider

$$u_{tt} = a u_{xx} + u_{yy} + b u_x + c u_y$$

The problem is strongly hyperbolic if $a > 0$ and weakly hyperbolic if $a = 0$ and there is catastrophic exponential growth if $a = 0$ and $b \neq 0$.

In the next section we will show that the ADM equations, linearized around flat, are only weakly hyperbolic for constant lapse function and shift vector. We can transform it into a strongly hyperbolic system if we choose the lapse function proportional to the trace of the solution. However, such a choice might introduce singularities.

Consider, for example,

$$u_{tt} = \alpha_x u_x + \alpha_{xx}.$$

If $\alpha = \alpha(x,t)$ is a given function, then the equation is weakly hyperbolic. If we choose $\alpha = u$, we obtain

$$u_{tt} = (u_x)^2 + u_{xx}.$$

Now the equation is strongly hyperbolic but we will, in general, encounter singularities due to the lower order term.

3 Second Order Initial Value Formulations For General Relativity

Our starting point are the ADM equations [2] for General Relativity. The 3-metric induced on the spacelike 3-surfaces $t = constant$ is denoted by γ_{ij},; all latin indices run over $1, 2, 3$.

From start we fix the shift vector equal to zero but keep a general lapse function α. Using the ADM equation for γ_{ij} to eliminate the extrinsic curvature from the other ADM equation we get a second order evolution equation for γ_{ij}

$$\partial_t^2 \gamma_{ij} = \alpha^2 \gamma^{lm} \Big(\partial_l \partial_m \gamma_{ij} + \partial_i \partial_j \gamma_{lm} - \partial_i \partial_l \gamma_{mj} - \partial_j \partial_l \gamma_{mi} \Big) + 2\alpha \partial_i \partial_j \alpha + \text{lot}, \tag{10}$$

where all derivatives are partial derivatives with respect to time and the coordinates on the $t = constant$ 3-surfaces. Here and below "lot" stands for "lower order terms", that is functions of α, γ_{ij} and their first derivatives. For the purpose of this paper, it is enough to say that all lower order terms are quadratic in first order derivatives.

We have to consider the constraint equations. The momentun constraint is

$$\gamma^{jl} \partial_t \Big(\partial_i \gamma_{jl} - \partial_j \gamma_{il} \Big) + \text{lot} = 0,$$

while the Hamiltonian constraint is

$$\gamma^{ij} \gamma^{lm} \Big(\partial_i \partial_l \gamma_{jm} - \partial_i \partial_j \gamma_{lm} \Big) + \text{lot} = 0.$$

We now linearize around a flat solution (Minkowski spacetime) in Cartesian coordinates, that is we make

$$\gamma_{ij} = \delta_{ij} + \varepsilon\, h_{ij} \qquad \text{and} \qquad \gamma^{lm} = \delta^{lm} - \varepsilon\, \delta^{lp} \delta^{mq} h_{pq} + \mathcal{O}(\varepsilon^2) \tag{11}$$

and keep terms linear in ε. Our new variable is h_{ij}. The constraint equations become

$$\partial_t \Big(\partial_i H - \delta^{lm} \partial_l h_{im} \Big) = 0$$

and

$$\delta^{ij} \delta^{lm} \partial_i \partial_l h_{im} - \Delta H = 0$$

where $H = \mathrm{tr}(h_{ij}) = \delta^{ij} h_{ij}$ and Δ is the Laplacian in $\mathbf{R}^3$. Both linearized constraint equations are satisfied if

$$\partial_i H - \delta^{lm} \partial_l h_{im} = 0. \tag{12}$$

Before linearizing (10) we have to make a choice of lapse. On the one hand the simplest possible choice $\alpha \equiv 1$ gives, after linearization,

$$\begin{aligned} \partial_t^2 h_{ij} &= \delta^{lm} \Big(\partial_l \partial_m h_{ij} + \partial_i \partial_j h_{lm} - \partial_i \partial_l h_{mj} - \partial_j \partial_l h_{mi} \Big) \\ &= \Delta h_{ij} + \partial_j \Big(\partial_i H - \delta^{lm} \partial_l h_{mi} \Big) - \partial_i \partial_j H \end{aligned}$$

so that (12) implies

$$\partial_t^2 h_{ij} = \Delta h_{ij} - \partial_i \partial_j H. \tag{13}$$

On the other hand, choosing

$$\alpha = \frac{k}{3} \mathrm{tr}(\gamma_{ij}) = \frac{k}{3} \delta^{lm} \gamma_{lm},$$

gives after linearization and using (12)

$$\partial_t^2 h_{ij} = k^2 \Big(\Delta h_{ij} - \frac{1}{3} \partial_i \partial_j H \Big). \tag{14}$$

We define $\boldsymbol{u} = (h_{11}, h_{22}, h_{33}, h_{12}, h_{13}, h_{23})^{\mathrm{t}}$ to analyze the hyperbolicity of equations (14) and (13).

Thus the matrix $P_0(\omega')$, as defined in Section 2.2, of the system asociated to (14) has constant eigenvalues: k^2 with multiplicity five, and $(2/3)k^2$ with multiplicity one. Also, this matrix has a uniformly linearly independent complete system of eigenvectors. Then according to Theorem 4, the system is strongly hyperbolic for any $k \neq 0$. $|k| \leq 1$ gives a system with characteristic speeds smaller or equal than one, while $|k| > 1$ would be an "unphysical" (though strongly hyperbolic) system with characteristic speeds higher than one.

The matrix $P_0(\omega')$ of the system asociated to (13) is uniformly diagonable but its eigenvalues are: 0 with multiplicity one and 1 with multiplicity five. Thus, according to Lemma 3, equation (13) is only weakly hyperbolic.

Another possibility to "cure" equation (13) is the usual choice of coupling the lapse function to the determinant of the 3-metric instead of coupling it to the trace as we have done above. This also leads to a strongly hyperbolic system.

4 Difference Approximations

Consider the simple model problem

$$u_t = u_x, \quad -\infty < x < \infty, \ t \geq 0 \tag{15}$$

with initial data

$$u(x,0) = f(x). \tag{16}$$

We are interested in solutions which are 2π-periodic in space. We want to solve the above problem by difference approximation. For that reason we introduce a gridlength $h = 2\pi/N$, N a natural number, gridpoints $x_\nu = \nu h$ and gridfunctions $u_\nu(t) = u(x_\nu, t)$. We also introduce the usual centered difference operators by

$$\begin{aligned} \partial/\partial x \quad &\sim D_0 u_\nu = (u_{\nu+1} - u_{\nu-1})/2h, \\ \partial^2/\partial x^2 &\sim D_+ D_- u_\nu = (u_{\nu+1} - 2u_\nu + u_{\nu-1})/h^2. \end{aligned}$$

Then we approximate (15),(16) by

$$(\tilde{u}_\nu)_t = D_0 \tilde{u}_\nu, \quad \nu = 0, 1, 2, \ldots, N-1 \tag{17}$$

with periodic boundary conditions

$$\tilde{u}_\nu(t) = \tilde{u}_{\nu+N}(t) \tag{18}$$

and initial conditions

$$\tilde{u}_\nu(0) = f_\nu. \tag{19}$$

(17) – (19) represents a system of ordinary differential equations which we solve with help of a standard ODE solver like the usual Runge-Kutta method.

We want to discuss the accuracy of the approximation. We assume that

$$f(x) = \sum_{\omega=-M}^{M} e^{i\omega x} \hat{f}(\omega), \quad M \leq N/2.$$

Then we can expand both the solutions of (15),(16) and (17)-(19) into Fourier polynomials

$$u(x,t) = \sum_{\omega=-M}^{M} e^{i\omega x} \hat{u}(\omega, t), \quad \tilde{u}_\nu(t) = \sum_{\omega=-M}^{M} e^{i\omega x_\nu} \hat{\tilde{u}}(\omega, t) \tag{20}$$

with

$$\hat{u}(\omega, 0) = \hat{\tilde{u}}(\omega, 0) = \hat{f}(\omega).$$

We introduce (20) into (15) and (17), respectively. Since

$$\partial e^{i\omega x}/\partial x = i\omega e^{i\omega x} \quad \text{and} \quad D_0 e^{i\omega x} = \frac{i \sin \omega h}{h} e^{i\omega x},$$

we obtain, for every frequency,

$$\hat{u}_t(\omega,t) = i\omega\hat{u}(\omega,t), \quad \hat{\tilde{u}}_t(\omega,t) = i\alpha\omega\hat{\tilde{u}}(\omega,t),\ \alpha = \frac{\sin\omega h}{\omega h}.$$

Therefore,

$$\hat{u}(\omega,t) = e^{i\omega t}\hat{f}(\omega), \quad \hat{\tilde{u}}(\omega,t) = e^{i\alpha\omega t}\hat{f}(\omega).$$

Thus there is a phase error

$$e = (1-\alpha)\omega t.$$

Also, the wave speed for the difference approximation depends on the frequency, i.e., there is dispersion for the difference approximation but not for the differential equation. This causes lots of difficulties if the solution is not properly resolved, i.e., there are not enough points/wavelength.

Instead of the above second order method, one can use higher order methods. This results in a remarkable improvement of the accuracy. In Table 1 we give the number of points/wavelength so that the numeric solution has a phase-error of 10% or 1% after calculating during q time periods with methods of different order.

Table 1. Points/wavelength

e	2nd Order	4th Order	6th Order
10%	$20\ q^{1/2}$	$7\ q^{1/4}$	$5\ q^{1/6}$
1%	$64\ q^{1/2}$	$13\ q^{1/4}$	$8\ q^{1/6}$

P. Huebner [3] and J. Thornburg [4], using fourth order accurate methods, have demonstrated the improved accuracy for the Einstein equations.

If we calculate with N points, then the solution consists of $M \sim N/2$ simple waves. Most of them have large phase errors. Therefore, the approximation is only useful if the Fourier expansion of the analytic solution converges rapidly. In particular, the part of the solution consisting of those waves with few points/wavelength travels in the wrong direction.

Consider

$$\hat{\tilde{u}}_{\nu t} = D_0\hat{\tilde{u}}_\nu$$

with highly oscillatory initial data

$$\hat{\tilde{u}}_\nu(0) = (-1)^\nu g_\nu, \quad g \text{ smooth function.}$$

Introduce a new variable by $\tilde{u}_\nu = (-1)^\nu w_\nu$. Then w solves

$$\begin{aligned} w_{\nu t} &= -D_0 w_\nu, \\ w_\nu(0) &= g_\nu, \end{aligned}$$

which approximates

$$\begin{aligned} w_t &= -w_x, \\ w(x,0) &= g(x). \end{aligned}$$

Thus $\hat{\tilde{u}}_\nu(t)$ represents a highly oscillatory wave which travels in the "wrong" direction. The usual way to control $\tilde{u}_\nu(t)$ is to add an artificial viscosity term.

As an example, consider

$$u_t = x\, u_x, \quad -1 \le x \le 1,$$

with boundary condition

$$u(-1,t) = -1, \quad u(1,t) = 1,$$

and the initial data

$$u(x,0) = -\cos\frac{\pi}{2}(x+1).$$

The solution forms an internal layer at $x = 0$ where the gradient becomes larger and larger. If we use the approximation

$$\tilde{u}_{\nu t} = x_\nu D_0 \tilde{u}_\nu,$$

then there will be a highly oscillatory wave traveling out of the layer.

We approximate the wave equation

$$u_{tt} = u_{xx}$$

by

$$\tilde{u}_{tt} = D_+ D_- \tilde{u}.$$

For the same level of accuracy, we need only half the number of points/wavelength. Also, there are no spurious waves which travel in the wrong direction.

5 Constraints

Using an example from fluid dynamics we want to demonstrate some of the problems which can arise when solving equations with constraints. Consider the Stokes problem

$$u_t + p_x = \Delta u, \tag{21}$$

$$v_t + p_y = \Delta v, \tag{22}$$

$$d =: u_x + v_y = 0, \tag{23}$$

in some domain $\Omega \times (0,T)$ with boundary conditions

$$u_n = 0 \quad \text{for} \quad (x,y) \in \partial\Omega, \qquad 0 \le t \le T.$$

Here u, v denote the velocity components, u_n the normal component and p the pressure. Differentiating the first equation with respect to x and the second with respect to y gives us, using the divergence relation $d = 0$,

$$\Delta p = 0. \tag{24}$$

We solve (21),(22) and (24) and think of (23) as the constraint.

One could be tempted to use, for p,

$$p = p_0, \quad (x, y) \in \partial\Omega, \quad p_0 \text{ given function}, \tag{25}$$

as boundary condition. However, then we would, in general, not preserve the constraint $d = 0$.

Differentiating (21) with respect to x and (22) with respect to y and using (24) gives us an equation for the divergence d,

$$d_t = \Delta d.$$

By assumption, $d = 0$ for $t = 0$ but we cannot guarantee that $d = 0$ at later times if we use the boundary conditions (25). We must use

$$d = 0 \quad \text{for} \quad (x, y) \in \partial\Omega, \ t \geq 0,$$

as boundary condition and we cannot give p.

Let $\Delta_h = D_{+x}D_{-x} + D_{+y}D_{-y}$. A typical difference approximation is given by

$$\tilde{u}_t + D_{0x}\tilde{p} = \Delta_h \tilde{u}, \tag{26}$$
$$\tilde{v}_t + D_{0y}\tilde{p} = \Delta_h \tilde{v}, \tag{27}$$
$$\Delta_h \tilde{p} \qquad = 0. \tag{28}$$

For the discrete divergence $d_h = D_{0x}\tilde{u} + D_{0y}\tilde{v}$ we then obtain

$$d_{ht} = -(D_{0x}^2 + D_{0y}^2)\tilde{p} + \Delta_h d_h, \tag{29}$$
$$\Delta_h \tilde{p} = 0. \tag{30}$$

The difficulty here is that

$$\Delta_h \neq D_{0x}^2 + D_{0y}^2$$

and therefore divergence is created. Instead of (28) one can use

$$\Delta_h \tilde{p} = \alpha d_h, \quad \alpha >> 1 \text{ constant}.$$

Then we can write (29) as

$$d_{ht} + (D_{0x}^2 + D_{0y}^2 - \Delta_h)\tilde{p} + \alpha d_h = \Delta_h d_h.$$

The damping term αd_h keeps d_h under control.

References

1. H.-O. Kreiss, J. Lorenz: 'Initial-Boundary Value Problems and the Navier-Stokes Equations'. In *Pure and Applied Mathematics, Vol. 136* (Academic Press Inc., San Diego 1989).
2. R. Arnowitt, S. Deser, C. W. Misner: 'The Dynamics of General Relativity'. In *Gravitation: An Introduction to Current Research*, ed. L. Witten (Wiley, New York 1962).
3. P. Huebner: Class. Quantum Grav. **16**, 2823 (1999).
4. J. Thornburg: gr-qc/9906022.